Mock Data Challenge

Robert Hatcher
Fermilab Computing Division

Cambridge Mtg. January 2004

Outline

- Purpose
- Data Sets
- New features of event generation and tracking
  - GMINOS geometry
  - Flux
  - Neugen3
  - Oscillations
- New Digitization procedure
  - PhotonTransport + DetSim
- Offline Reconstruction

Purpose

- Produce files to exercise the full analysis
  - test of overall framework and data handling
  - test both NearDet and FarDet reco chains
- Practice the extraction of physics signal
  - conventional δm^2_{23} signal
  - possible $_{e}$ and $_{sterile}$ signals
- Possibly Atmospheric/Up-Down muons (no veto shield)
  - useful to have single common set in tape robot, generated/processed consistent with above set
- *Not* a test of calibration chain

Data Sets

- Challenge set
  - oscillations (far), modified flux, modified Neugen3
  - truth information removed (from public access)
  - FarDet: 3yr at 2.4e13 pots/spill
  - NearDet: overlayed, 25 x FarDet statistics in target region
    ~100k fid events + 1 week's horn off running (302400 spills?)
- Monte Carlo set
  - no oscillations (far), nominal flux, nominal Neugen3
  - truth information retained
  - FarDet: 10x 'data' POTs (~100k events)
  - NearDet: 1x 'data'
- Near statistics will be hard to achieve
  - flux reweighting slows generation down
  - rock event generation is very time intensive
    - may not need to be independent; possible to reuse these events?
  - overlay processing not entirely insignificant
- Reconstruction processing will be a major effort

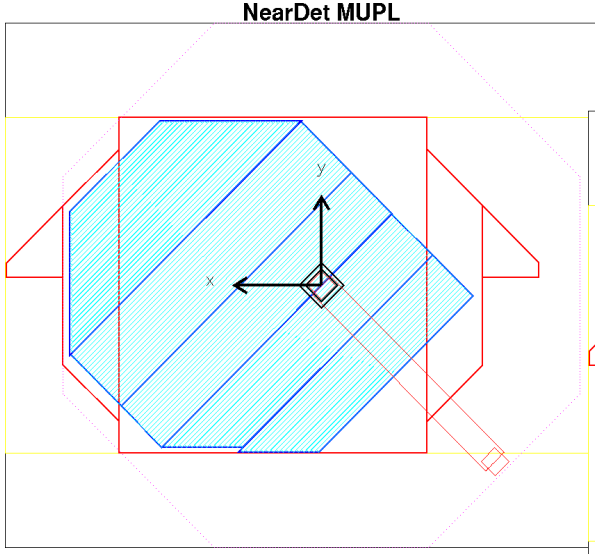
Rough Time Estimates

- FarDet example file:
  - 3.8e21 POTs = 36538 events = 41.9 GHz*hrs
  - this is ~ size of MC data set
  - estimate ~13 days to reconstruct (@30sec/evt)
- NearDet example file:
  - 1000 spills (8m,20s beam time) at 4e13 POTs/spill assembled from
    - 42016 events in the NearDet (40.6 GHz*hrs)
    - 20877 events generated in the surrounding rock that left energy in a scintillator strip (135.1 GHz*hrs)
    - 9.6 GHz*hrs overlay procedure (and redigitization)
  - a target cylinder R<50cm, 20<ipln<59 is ~0.65% of total mass means that only 1 in 3.8 spills (at this higher POTs/spill) has an event in this volume; even worse if one requires 25cm radius
- Near challenge, MC sets
  - ~1.4e7 total events or ~5.75e5 spills (13.3days beam time)
  - ~113.4 GHz*days + ~377.5 GHz*days = 16.4 GHz*months each
  - tricky file management issues due to overlaying
  - ? estimate of reco time
-

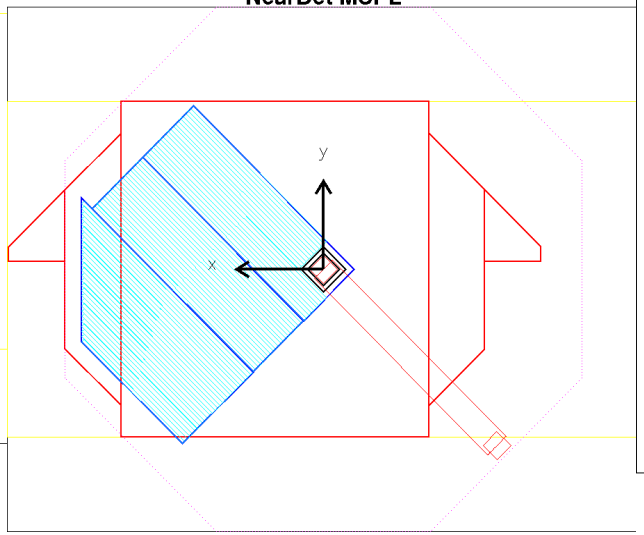
GMINOS Geometry

- Module based planes (right # strips/module, NearDet not trimmed down larger octogon)
  - official strip, WLS pigtail lengths
  - average clear fiber lengths (for each side of module type and view)
  - 0.2mm gaps between strips and 2mm between modules
  - (as before) modules include aluminum skin, but no extra crimp
  - bypass regions still not inactive (i.e. strips go up to the coil collar)
- Detector hall sizes adjusted to *as built*
- FarDet plane spacing 5.949/5.943 vs. nominal 5.94cm
- FarDet coil is a foam of average material
- New DearDet steel shape

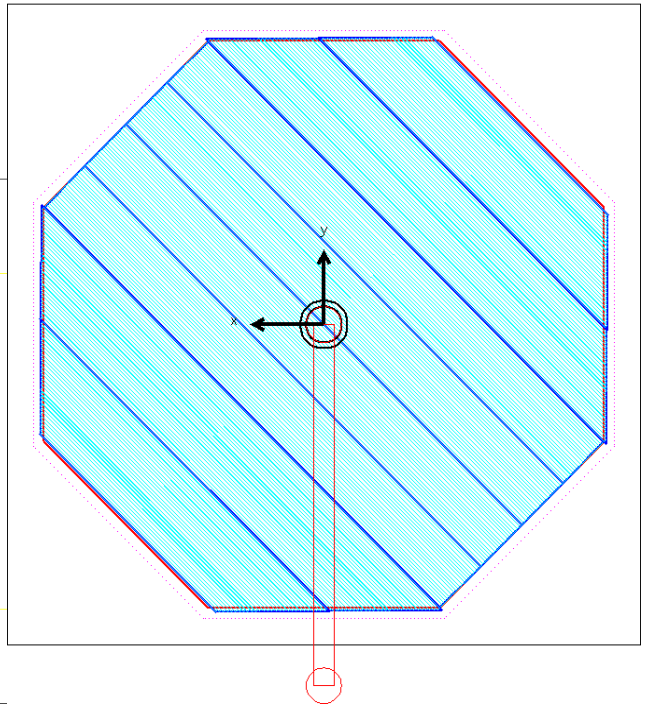
NearDet MUPL



NearDet MSPL

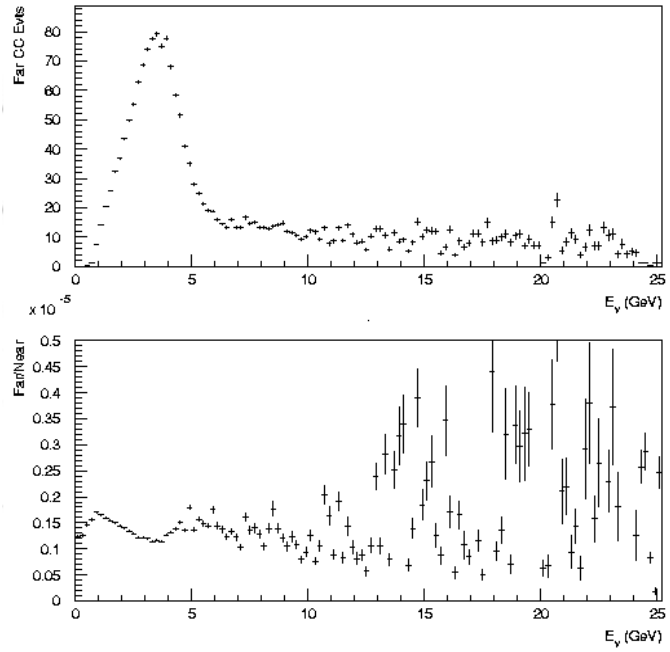


FarDet M2PL

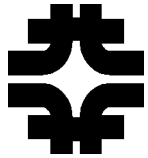


Flux

- GMINOS directly reads column-wise ntuples (flux method 5)
  - currently gNuMI version v16 (vacuum window mod in v15, no change 15 to 16?)
  - Mark expects to have v17 RealSoonNow
- Reweight energy and probability based on position on detector face
- Challenge set will have additional reweighting of hadron production model
  - wgt ba[illegible]



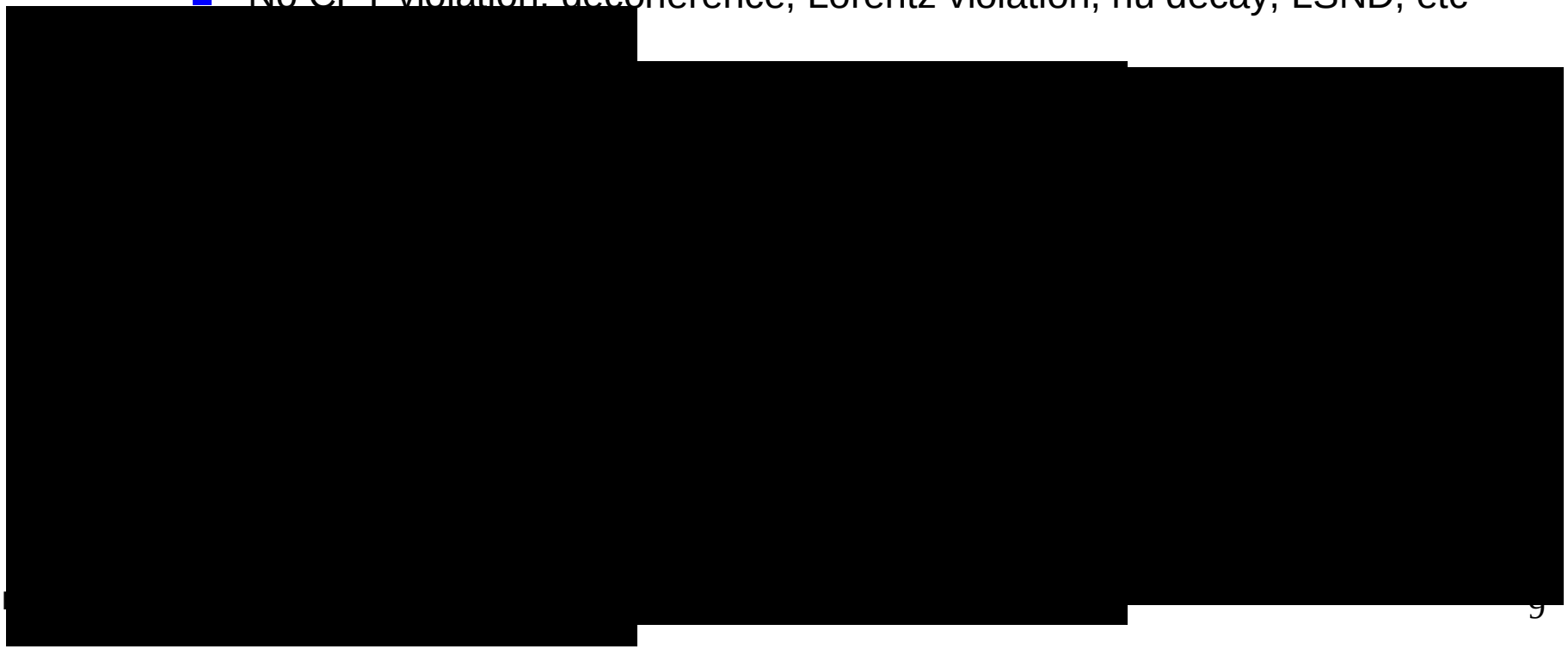
Neugen3 (Hugh Gallagher)



- Rewrite of older NEUGEN
  - Improved internuclear scattering, cross sections, etc...
- Challenge data will use same set of physics models, but with tweaked parameters.
  - QE mass (`gv_ma_qel`): 1.032 ± 4%
  - Resonance mass (`gv_ma_res`): 1.032 ± 4%
  - DIS fudge factor (`gv_mdc_fudge`): 1.0 ± 3%
  - in principle modified values are a black box until opened at a later date (held as parameters in `physconst.inc`)

Oscillations (ala Mark Messier)

- 100% chance of $\nu_\mu \in \nu_\tau$ oscillations
  - $\log_{10}(\Delta m^2_{23})$ gaussian(SK,1.5*1σ_{SK})
  - $\sin^2(2\theta_{23})$ gaussian(SK,1.5*1σ_{SK}) restricted to physical region
- 75% chance of $\nu_\mu \not\subset \nu_e$
  - parameters uniformly up to CHOOZ 95% CL limit
- 75% chance of sterile oscillation
  - parameters uniform up to SK 95% CL limit
- No CPT violation, decoherence, Lorentz violation, nu decay, LSND, etc



Digitization

- Digitization is the process of turning truth information about energy deposition into DAQ-like data output
- Nathaniel Tagg has written two very substantial packages
  - PhotonTransport
    - dE into scint photons, capture by WLS, transport to end of WLS
    - three models with differing levels of sophistications (vs. time)
    - can account for observed larger effective average index of refraction by correctly modelling tails due to spiral rays
  - DetSim
    - convert photons at cathode into RawDigitDataBlock
    - account for PMT effects, electronics (VA,QIE), triggering
  - Two packages replace RerootToRawDataModule (and F77 code)
    - see PhotonTransport/macros/simple.C for example script
    - see PhotonTransport/doc and DetSim/doc for details

Offline Reconstruction

- Process MC and Challenge sets with a frozen release of the code
- Do it at FNAL on the Batch Farm to avoid duplication of effort and possible inconsistencies in reconstruction code versions
- Make the fully processed MC set available from tape store
- Make the fully processed and truth-stripped Challenge set available
  - keep private copy of challenge set truth information to be revealed at a later date (detailed post mortem)

Issues

- GEANT hadronic model: GHEISHA, FLUKA?
- distribution of data sets to non-FNAL sites: tapes, DVDs?
  - ntuples probably doable; candidate files?
  - who want it? who does the media copying?
- known and unknown deficiencies in 'standard' reconstruction
  - is near reco ready to run?
  - outstanding Heisenbug
  - recent Cand change losing handles?
- alternative reconstruction modules?
  - media copying vs. extra cycles
  - combinatorics headaches
  -